 ICCES, vol.10, no.1, pp.3-4

Segmentation methods for human motion analysis from image sequences

Maria João M. Vasconcelos[1], João Manuel R. S. Tavares[1]

Summary

In the last years, researchers from the Computational Vision working field have been developing new methods to perform image segmentation for human motion analysis. The development of computational techniques suitable to automatically identify the structures involved is necessary to obtain more representative and robust features to be further used in the analysis of human motion from image sequences.

The first step of human motion analysis from image sequences is strongly related with image segmentation. In fact, the first goal of any system designed for this aim is the identification of the structures' features to be analysed in the image frames. If, for a human, the task of identifying moving structures in images is more or less trivial, computationally this task has proofed not to be so simple.

Image segmentation methods need to deal with some challenges concerning the image sequences which are used. As examples of these challenges we can mention: lighting conditions that can change along the sequences; occlusion problems, when the structure does not remain inside the workspace or when the structure is partially occluded; the existence of dynamic backgrounds, when the camera is in motion or the scene is changing; or multiple moving structures, when there is more than one structure moving in the workspace at the same time. It is not straightforward to develop methods suitable to deal with all these problems and difficulties at once, so it is common to make some assumptions and simplifications; however, each day more and more robust and accurate methods are being developed.

The use of an edge detection algorithm, by itself, is obviously not enough to identify a structure in an image sequence. The most typical method of image segmentation is background subtraction, which involves the calculi of a reference image followed by the subtraction of each frame of the image sequence from the reference and further threshold of the result. The simplest form is using a time-averaged background image as reference but it requires a training period absent of foreground objects. Other possibility is describing each pixel in the scene by a mixture of Gaussian distributions, where the weight parameters of the mixture represent the time proportions that those colours stay in the scene, so background components will be the ones with the highest probable colours. However, this last method usually fails in busy environments where a clean background is rare.

[1]Institute of Mechanical Engineering and Industrial Management / Faculty of Engineering, University of Porto, Porto, Portugal. Email: {maria.vasconcelos, tavares}@fe.up.pt

Other possibility is the use of a method based on Bayes decision theory to detect foreground structures from complex image sequences. In the first step of this method, non-change pixels in the image stream are filtered out by simple background and temporal differences. Then, the detected changes are separated as pixels belonging to stationary and moving structure according to inter-frame changes. After that, the pixels associated with stationary or moving structures are classified as background or foreground based on the learned statistics of colours through the use of the Bayes decision rule. Finally, foreground structures are segmented by fusing the results from both stationary and motion pixels and the background model is updated. This method showed to work well in complex backgrounds including sequences with variable light conditions and shadows of moving structures.

In this work we will explore in more detail the two segmentation methods referred, present some experimental results and address possible practical applications related with human motion.